of folding-back and averaging-out were repeatedly applied until the expected values of benefit and cost are eventually calculated at the first decision-node. **RESULTS:** Comparing Teb to Tc, we obtained the incremental cost and benefit for $(p(1-\beta) + \alpha(1-p))(Ca - Cb)$ and $p(1-\beta)(Ea - Eb)$, respectively. Then, dividing the former by the latter resulted in the formula: r-ICER = (1+k) x ICER, where $k = (\alpha(1-p))/((1-\beta)p)$. It implies that the risk-adjustment corrects the underestimation of ICER since k takes a positive value. Hence, the efficiency frontier defined by a series of r-ICERs transformed into an inferior position, compared to the original one. **CONCLUSIONS:** The r-ICER can correct an underestimate of the standard ICER and will be useful in risk-sensitive evaluation using ICERs including the efficiency frontier.

**PMC2**

## A STANDARDIZED EVIDENCE BASED APPROACH TO ASSESS NON-TRADITIONAL OUTCOME MEASURES FOR USE IN HEALTH CARE DECISION MAKING: THE DIABETES EXAMPLE

Badia X[1], Vieta A[1], Sacristán JA[2]
[1]HEOR IMS Health, Barcelona, Spain, [2]Clinical Research Department Lilly, Madrid, Spain

**OBJECTIVES:** Non-traditional outcomes (NTOs), those related with patient reported outcomes (PROs), economic and non-traditional clinical outcomes, are frequently being used to assess health interventions. We propose a standardized approach to assess the utility of NTO measures for use in health care decision making. **METHODS:** A systematic review of NTOs in Type 2 Diabetes Mellitus (T2DM) was conducted. Inclusion and exclusion criteria, data sources, search strategy, and data extraction and quality assessment of the studies and NTOs were defined. The degree of recommendation of each NTOs was based on the quality of the outcome and the scientific evidence to support it. Two independent reviewers carried out each activity. **RESULTS:** NTOs' were assessed within a three-grade quality scale in terms of feasibility, validity, sensitivity, reliability, comparability and understanding. NTOs were categorized as key, important, and with not enough evidence to support its use in health decision-making. Case study in T2DM: 3805 citations and 235 potentially eligible full articles were retrieved and 153 studies met the inclusion criteria. Eighty-eight (5 clinical, 54 humanistic and 29 economic) NTOs in T2DM were retrieved. A total of 21.6% of the NTOs were considered key, 36.4% important and for 42% not enough evidence was found to support its use in T2DM. **CONCLUSIONS:** An evidence based understanding of NTOs' validity to measure treatment outcomes in different conditions is needed since clinicians and payers may use them for decision-making purposes.

**PMC3**

## CRITICAL REVIEW OF PUBLISHED COST EFFECTIVENESS ANALYSES OF ORAL TRIPTANS IN THE MANAGEMENT OF ACUTE MIGRAINE: DIFFERENT MEASURES OF EFFICACY EXTRACTED FROM THE SAME META-ANALYSIS

Peura P, Purmonen T, Martikainen JA, Turunen JH
University of Kuopio, Kuopio, Finland

**OBJECTIVES:** Meta-analyses are recommended and frequently utilized in cost-effectiveness analyses. The aim of this study was to assess if efficacy data obtained from one meta-analysis is used differently in various cost-effectiveness analyses of oral anti-migraine medications (triptans). **METHODS:** A literature review was conducted in order to identify studies, assessing the cost-effectiveness of triptans, that used the meta-analysis by Ferrari et al. (2002) as the primary source of efficacy. Studies were included in the literature review if they assessed the cost-effectiveness of two or more triptans (almotriptan, eletriptan, frovatriptan, naratriptan, sumatriptan, rizatriptan or zolmitriptan) in the adult population using an average or incremental cost-effectiveness ratio as an outcome measure. **RESULTS:** In total, 12 studies were included in the final review. Five different measures of efficacy were used in the reviewed studies. These were "sustained pain free with no adverse effects" (n = 6), "24 hour sustained response" (n = 3), "sustained pain free" (n = 4), "pain free at two hours" (n = 1), and "quality adjusted life year" (QALY) (n = 2). In addition, four alternative methods were used to calculate the "sustained pain free with no adverse effects". In total, this resulted in eight alternative measures of efficacy derived from a single meta-analysis. **CONCLUSIONS:** The choice of efficacy measure can differ between different cost-effectiveness studies, even those based on same literature source. This may affect the applicability of the cost-effectiveness analysis in decision making. However, the underlying reasons for the variation in the choice of efficacy measure, and the association between choice of efficacy measure and the study conclusion need thorough research. Therefore, there is still room for conventional one-way sensitivity analyses that enable evaluation of the effects of "qualitative" data source uncertainties on the study results.

**PMC4**

## LARGELY IGNORED: THE IMPACT OF THE THRESHOLD VALUE FOR A QALY ON THE IMPORTANCE OF A TRANSFERABILITY FACTOR

Vemer P, Rutten-van Molken M
Erasmus MC, Rotterdam, The Netherlands

Recently, several checklists have been developed which systematically check the transferability of cost-effectiveness (CE) studies between jurisdictions. Some even provide a quantitative score of the degree of transferability. The interpretation of such a score is, however, difficult. In addition, the threshold value for a QALY is a factor that has been given little consideration in these checklists. The importance of a factor as a cause of between-country differences in CE depends on this threshold: factors that explain most of the differences in CE at a low threshold need not be the same factors at a higher threshold. **OBJECTIVES:** To compare the impact of nine potential causes of variation in CE of smoking cessation support (SCS), at different thresholds for the willingness-to-pay per QALY (WTP). **METHODS:** A model-based study compared the cost-effectiveness of SCS between six Western-European countries. For several values of WTP, we investigated the impact of between-country differences in nine factors on the incremental net monetary benefit (INMB). The factors were demography, smoking prevalence, mortality, epidemiology and costs of smoking-related diseases, resource use and unit costs of SCS, utility weights and discount rates. **RESULTS:** Currently, SCS is not reimbursed in The Netherlands, corresponding to a WTP of € 0. With a WTP below €1000, the factors most responsible for between-country differences in INMB are resource use and unit costs of SCS and the costs of smoking-related diseases. Utility values have little impact. At a threshold above €10,000, between-country differences are primarily due to different discount rates, utility weights and epidemiology of smoking-related diseases (incidence and mortality). Costs of smoking-related diseases have little impact above €20,000. At all thresholds, demography has little impact. **CONCLUSIONS:** When judging the transferability of a CE study to another jurisdiction, we should consider the between-country differences in threshold values per QALY.

**PMC5**

## MEASURING AND MONITORING THE REAL-WORLD COST-EFFECTIVENESS OF NEW TECHNOLOGIES IN HOSPITALS

Neurohr C, Welte R
GlaxoSmithKline, Munich, Germany

**OBJECTIVES:** Due to setting specific characteristics the economic attractiveness of new technologies might vary between hospitals. We compared classical and statistical process control (SPC) methods for measuring and monitoring the real-world cost-effectiveness of new technologies in hospitals. **METHODS:** A systematic literature review was performed in PubMed in April 2009 to identify studies applying classical and SPC methods for investigating the (cost-)effectiveness of changes in inpatient health care processes. Both methodologies were compared using a predefined set of criteria such as *accuracy, flexibility, informative value, suitability and user-friendliness*. **RESULTS:** Classical statistical methods based on 'time static' (cross-sectional) statistical analysis with aggregated data are widely used. With the ability to detect statistical significant differences classical methods may provide higher accuracy. They are characterized by large one-time data collections to evaluate the impact of a process change for a pre-specified time period, limiting their flexibility. SPC methods which analyze time series data by monitoring a process over time have been used rarely but their application is increasing. They combine time series analysis with a graphical representation of the data. Patterns in time series data contain important information which other methods reliant on averages (or other summary statistics) could mask. By providing continuous feedback SPC is capable not only of detecting the results of process changes earlier but also of monitoring the process sustainability. SPC can be applied to routine data easier as it is typically less sensitive to statistical issues. Furthermore, the graphical representation of the data has advantages because statistical measures such as P-values are often poorly understood and misinterpreted. **CONCLUSIONS:** Both methodologies are suitable for measuring the (cost)-effectiveness of changes in health care processes. SPC seems to be the preferred methodology under real-world conditions to support decision-making although it commonly does not achieve the accuracy of classical statistical methods.

**PMC6**

## THE NATURE AND SCALE OF INADEQUATE REPORTING OF CONTINUOUS OUTCOMES FROM FOUR SYSTEMATIC REVIEWS

Singh N, Dimri S, Mahajan A, Agrawal R, Ubhadiya BS, Narvilkar P, Bhanderi M
Heron Health Private Ltd, Chandigarh, India

**OBJECTIVES:** Inadequate reporting of continuous outcomes is a major problem while performing meta-analysis. The objective of this study is to estimate the nature and scale of inadequate reporting of continuous outcomes. **METHODS:** Reporting quality of continuous outcomes (baseline, endpoint and change from baseline data) was analysed across four disease area reviews conducted in lipid disorder, overactive bladder, multiple sclerosis and rheumatoid arthritis. Reporting quality was considered inadequate when either number of patients analysed (N) was not reported or error term was missing/could not be calculated from the reported statistics. Analyses were conducted using STATA 9.2. **RESULTS:** In total 12,236 reported outcomes across the four systematic reviews were included in the analysis. Inadequate reporting of continuous outcomes was frequent and observed for 3912 (31.97%) outcomes. The estimate of inadequate reporting varied across selected reviews and ranged from 27.10% to 45.77% in lipid disorder and overactive bladder, respectively. Of the outcomes reported inadequately, number of patients analysed (N) was missing for 8.82% whereas error term was missing or could not be calculated for 91.18% outcomes. When the reporting quality was analysed by the outcome type, it was observed that change from baseline data were often reported inadequately (49.63%) compared to baseline & endpoint data (25.00% and 24.51%, respectively). **CONCLUSIONS:** Inadequate reporting of continuous outcome was frequently observed among the selected reviews especially for the change from baseline outcomes. The results demonstrated that for majority of the outcomes, the error term was either missing or could not be calculated from the given statistics. This inadequacy of reporting could have a significant impact on the results of meta-analysis. Our results are indicative of outcome reporting bias which needs to be investigated further.